0022-1767/93/1519-4732$02.00/0
The Journal of Immunology
Copyright © 1993 by The American Association of Immunologists

Vol 151, 4732–4742, No. 9, November 1, 1993
Printed in U.S.A.

Self-Peptides from Four HLA-DR Alleles Share Hydrophobic Anchor Residues Near the NH_2-Terminal Including Proline as a Stop Signal for Trimming[1]

Harald Kropshofer,* Heiner Max,* Thomas Halder,* Matthias Kalbus,* Claudia A. Muller,† and Hubert Kalbacher[2]*

*Center of Medical and Natural Sciences, University of Tübingen, †Medical University Clinic, University of Tübingen, W-7400 Tübingen, Germany

ABSTRACT. Naturally processed MHC class II-associated peptides proved to be heterogeneous in size, varying from 13 to 25 amino acids. Truncation variants suggested sequence motifs that afford the amino termini to be shifted for obtaining an alignment: a 9- to 11-residue core region that is bordered by primary anchor residues is surrounded by extra sequences of variable lengths and hitherto unknown functions. Herein we present bulk sequencing analyses of self-peptides from four HLA-DR alleles and HLA-DQw7 clearly showing that the length of most of the NH_2-terminal preanchor sequence is limited to 1 to 3 residues. Most strikingly, proline is the dominant residue reappearing at positions 2 and 3 in any allele. Proline revealed to function as a stop signal for NH_2-terminal trimming as well as a secondary anchor: crude cytosolic and endosomal peptide fractions could be processed by aminopeptidases in vitro, whereupon DR1 binding peptides with increased affinity were generated. In addition, aminopeptidase treatment of DR1:self-peptide complexes implied that proline together with sterical constraints of the MHC molecule do protect the peptides' NH_2-termini from further processing, whereas their COOH-termini were accessible to cathepsin B processing. Finally, bulk sequencing profiles contained signals from further putative anchor residues clustering in the NH_2-terminal region: tyrosine, phenylalanine, leucine, isoleucine, and valine are enriched at positions 2 to 4 in DR1, DR5, and DR6, however, at positions 4 to 6 in DR3. Isotype-specificity is demonstrated by DQw7 displaying glutamine and asparagine at position 2. Obviously, the degenerate occurrence of aromatic or aliphatic side chains close to the NH_2-terminal guarantees for essential interactions with a hydrophobic pocket of the investigated DR molecules. Most probably, this pocket is located in the nonpolymorphic DR α-chain rationalizing previous findings of promiscuous peptide binding to different DR alleles. *Journal of Immunology,* 1993, 151: 4732.

Peptide fragments derived from foreign or self-protein Ag are constitutively presented on the surface of an APC tightly associated to molecules of the MHC (1). Because the availability, sequence, and length of certain peptides appear to determine whether normal, alloreactive, or autoaggressive T cell clones are activated or eliminated, a huge amount of work is focused on the elucidation of peptide motifs relevant for binding to MHC molecules. Sequence analyses by Edman degradation (2–4) and mass spectroscopy (5) revealed allele-specific binding motifs for MHC class I-associated self-peptides, ranging in

Received for publication February 9, 1993. Accepted for publication July 28, 1993.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] This work was supported by the Sonderforschungsbereiche 120 and 323. H.K. was supported by the Deutsche Forschungsgemeinschaft through Grant 767/1. H.K. was supported by the University of Tübingen through Grant IV 1.3 KJH/91.

[2] Address correspondence and reprint requests to Dr. Hubert Kalbacher, Center of Medical and Natural Sciences, University of Tübingen, Obere Im Himmelreich 7, 72074 Tübingen, Germany.

size from 8 to 10 amino acids. Similar approaches applied to naturally processed peptides from murine MHC class II could prove that these peptides were longer and rather heterogeneous in length (13 to 18 amino acids), also displaying allele-specific anchor residues at certain positions (6–8). The only human class II alleles that have been investigated for self-peptide motifs so far are HLA-DR1 and -DR11. The first study on DR1[3] described a 16-mer peptide with a hydrophobic two-residue contact motif being crucial for binding (9). Most recently, this finding could be largely confirmed by use of M13 epitope libraries (10). An alternative sequencing study with DR1 reported peptides varying from 13 to 25 residues, the putative motif tolerating a positively charged as well as a bulky hydrophobic side chain near the NH_2-terminal (11). The potential peptide binding motif for HLA-DR11/Dw52 was suggested to consist of an aromatic amino acid followed by a basic amino acid five positions after the aromatic residue and a second basic amino acid seven positions after the aromatic anchor (12). General agreement exists about the first anchor residue of class II-associated peptides having variable distance from the NH_2-terminal (8, 9, 11). That is the reason why the pooled peptide sequencing method, previously applied to deduce class I peptide motifs (3), is thought to be difficult to apply to class II self-peptide mixtures (8, 11). However, most naturally processed peptides revealed rather short NH_2-terminal preanchor sequences (8, 9, 12). Therefore, bulk sequencing should allow to prove the existence and the chemical nature of putative anchor residues, especially in the region close to the NH_2-terminal of the peptides.

Applying this technique in this study, we succeeded in verifying the isotype-specific occurrence of hydrophobic anchor residues near the NH_2-termini of pooled self-peptides from four HLA-DR alleles. Most significant, proline occurring frequently at positions 2 and 3 in the sequencing profiles of any allele revealed to be important for peptide binding as well as for NH_2-terminal processing: treatment of cytosolic and endosomal peptides by aminopeptidases in vitro made them highly potent competitors in a binding assay with influenza matrix peptide IM (19–31). This is probably due to proline functioning as a stop signal for aminopeptidases. Our sequencing and binding data are fully consistent with the observation that self-peptides bound to HLA-DR1 are resistant to aminopeptidase treatment but are further COOH-terminally truncated by cathepsin B in vitro reemphasizing the protection of the peptides' NH_2-termini and the susceptibility of their COOH-termini.

Materials and Methods

Cells

The EBV-transformed homozygous cell lines WT-100 (DR1Dw1), COX (DR17Dw3 ≡ DR3), KRO (DR11Dw5 ≡ DR5), and AMALA (DR14Dw16 ≡ DR6, DQw7) were used as a source for the isolation of HLA-DR and -DQ molecules. The cell lines were maintained in vitro at 37°C, 5% CO_2 by culture in RPMI 1640 medium (GIBCO, Grand Island, NY), containing 5% heat-inactivated FCS (GIBCO), 2 mM glutamine, and antibiotics (GIBCO).

Isolation of HLA-DR and -DQw7

Cells were lysed and the resulting homogenate prepared for affinity chromatography with mAb L243 (anti-DR) (13) and TÜ22 (anti-DQ) (14), essentially as described elsewhere (15).

Preparation of HLA-associated self-peptides

Naturally processed peptides were released by adjusting a solution of 500 μg of solubilized HLA-DR or -DQ to pH 2.2 by addition of trifluoroacetic acid. After a 3-h incubation at 37°C, released peptides were separated by ultrafiltration with a Microsep 10 microconcentrator (Filtron, Karlstein, Germany). The flow-through provided the acid-eluted peptide pool, which was concentrated to 100 μl in a speed vac. Peptides were freed from salt and detergent by microbore reverse-phase HPLC. We used an Aquapore OD-300 column (250 × 1.0 mm, 7 μm; Brownlee) (Applied Biosystems, San Jose, CA) and a Merck HPLC system, (Merck-Hitachi, Dornstadt, Germany) as previously described (15). Buffer A, 0.06% TFA, H_2O; buffer B, 0.052% TFA, 80% acetonitrile. Gradient: 0 to 120 min, 0 to 80% B, 120 to 130 min, 80 to 100% B. Material eluting with 10 to 35% acetonitrile was pooled and dried in a speed vac for sequence analyses.

Preparation of cytosolic peptides

The homogenate from 2 × 10^8 WT-100 cells was prepared as for affinity chromatography (see above). Peptides were separated by ultrafiltration with a Microsep 10 microconcentrator (Filtron). The flow-through was lyophilized and freed from salt and detergent by microbore reverse-phase HPLC as described above.

Preparation of endosomal peptides by Percoll density gradient fractionation

Endosomal fractions were prepared according to a standard protocol with the following modifications: WT-100 cells (10^9) were suspended in 10 ml of 10 mM Tris-HCl buffer,

[3] Abbreviations used in this paper: DR1, DR3, DR5, DR6, and DQw7 are human MHC class II molecules of the DR and DQ haplotype; AMCA, the fluorophor 7-amino-4-methylcoumarin-3-acetic acid; (AMCA)-DR1-33(6–16), an NH_2-terminally labeled DR1-associated self-peptide from amino acid position 6–16; DR1-33(6–16)-(AMCA), a COOH-terminally labeled DR1-associated self-peptide from amino acid position 6–16; AAP, alanyl-aminopeptidase; LAP, leucyl-aminopeptidase; Ii, invariant chain; HPSEC, high-performance size-exclusion chromatography.

pH 7.8, containing 0.2 mM PMSF (Sigma Chemical Co., St. Louis), 5 μM leupeptin, 10 μM pepstatin, and 1 μM chymostatin (Boehringer, Mannheim, Germany). Lysis was achieved using a Teflon homogenizer with a 10-μm gap. Nuclei and intact cells were removed by centrifugation (1000 × g, 10 min). The supernatant was mixed with a concentrated Percoll suspension (Sigma Chemical Co.) and a solution of 2.5 M sucrose, 10 mM Tris-HCl, pH 7.8, to a final volume of 60 ml, adjusting a concentration of 20% Percoll. A continuous Percoll gradient was generated by centrifugation of this suspension (60,000 × g, 50 min). Gradient fractions were drawn and endosomal fractions corresponding to a density of 1.030 – 1.045 were separated. They were diluted 10-fold by addition of 10 mM Tris-HCl, pH 7.8, 250 mM sucrose and centrifuged in a Beckman SW 28 swing-out rotor (100,000 × g, 120 min). Pellets were lysed in 10 mM Tris-HCl, pH 7.8, 1.0% (wt/vol) Triton X-100 for 10 min. Endosomal peptides were obtained by ultrafiltration and reverse-phase HPLC as described above. The endosomal origin of the described fractions could be verified by the accumulation of transferrin after pulsing whole cells with fluorescently labeled transferrin.

Treatment of peptide fractions with LAP

Peptide fractions were normalized by their OD at 220 nm before use. Digestions with LAP (Boehringer, Mannheim) were performed in 40 μl of 100 mM sodium phosphate buffer, pH 8.0, 0.1% (wt/vol) Zwittergent-12 (Calbiochem, San Diego, CA) for 16 h at 37°C using 25 ng/μl of the enzyme and 0.3 to 1.6 μg of peptide.

Treatment of DR1:peptide complexes by AAP and cathepsin B

The 100 μg of DR1 solubilized in 200 μl 100 mM sodium phosphate, pH 7.0, 0.1% Zwittergent-12 (Calbiochem) were coincubated with 1 μg AAP (Boehringer, Mannheim) for 16 h at 37°C and with 0.1 μg cathepsin B (Calbiochem) in 200 μl 100 mM sodium phosphate, 8 mM DDT, pH 6.0, for 16 h at room temperature, respectively. Peptides were released and separated as described above. In parallel, samples containing AAP or cathepsin B were supplemented with AMCA fluorescently labeled DR1-associated self-peptides DR1–33(6–16), ILLSKKHLNK(-AMCA)I in one letter code, or DR1–33(6–14), (AMCA)-ILLSKKHLN, respectively. Both peptides were digested for 30 min under the same conditions as described above. Each sample was analyzed by reverse-phase-HPLC on a Nucleosil C4 column (125 × 4 mm, 5 μm, Grom, Tübingen, Germany) with fluorescence detection (excitation wavelength = 350 nm, emission wavelength = 450 nm). Gradient: 0 to 30 min, 5 to 95% B. Buffer A and B were th same as described above.

Sequence analyses

Edman degradation was performed in a pulsed liquid sequencer 477A equipped with an on-line PTH-amino acid analyzer 120 A (Applied Biosystems Foster City, CA) essentially as previously described (3, 9). All samples were normalized to a total yield of 500 to 1000 pmol in the starting cycle. Peptide pools of each allele were sequenced in duplicate or triplicate.

Peptide synthesis

IM(19–31) was synthesized by continuous flow solid-phase peptide synthesis using a MilliGen (Eschborn, Germany) 9050 synthesizer based on Fmoc/Bu^t strategy. The peptide was purified by HPLC and the identity tested by ion spray mass spectrometry.

Fluorescence labeling

NH_2-terminal labeling was performed as previously described (15).

HPSEC peptide binding assay

Solubilized HLA-DR1 (2 μg) was coincubated with NH_2-terminally AMCA labeled peptide (AMCA)-IM(19–31) (0.1 μg) in 50 mM sodium phosphate, pH 5.0, 0.1% (wt/vol) Zwittergent-12, supplemented with a cocktail of protease inhibitors (see above) for 48 h at 37°C. As competitors aliquots of different peptide fractions, previously digested by aminopeptidases or not, were added in 3- to 16-fold molar excess relative to (AMCA)-IM(19–31), as judged by the OD at 220 nm. All samples were analyzed on a Pharmacia (Piscataway, NJ) Superdex 75 HR 10/30 high performance gel filtration column essentially as described elsewhere (16).

Results

Acid release of naturally processed self-peptides

MHC class II molecules from four EBV-B cell lines, each of them homozygous for the allele of interest, were purified by immunoaffinity chromatography and checked for purity by SDS-PAGE and high performance gel filtration (data not shown). Unspecifically bound low-molecular-weight material was separated by ultrafiltration over an Amicon YM 30 membrane. High affinity bound peptides were released by treatment with TFA before ultrafiltration over a Microsep 10 membrane. The material in the flow-through was then fractionated by microbore reverse-phase HPLC on a C18 column.

We obtained an allele-specific pattern of peaks for each allele, uniformly eluting between 10 and 35% acetonitrile, as shown for HLA-DR1 (Fig. 3). As a control, we used glycine coupled to Sepharose. We treated this column and

the eluate accordingly. The material we obtained from this mock precipitation did not show any prominent peaks eluting with 10 to 35% acetonitrile in the reverse-phase HPLC separation (data not shown). It was collected and taken for pool sequencing as well.

Sequence analyses of pooled peptides

Based on the assumption that certain amino acids crucial for anchoring naturally processed peptides to the relevant MHC molecule do reoccur at defined positions, those residues should accumulate in the corresponding sequencing cycles. With the Edman degradation technique sequentially cleaving off residues from the NH_2-terminal of a peptide, preanchor sequences of highly variable lengths (0 to 9 residues), as found for single peptides derived from HLA-DR1 (11), should make this approach useless for defining anchor residues. However, recent findings concerning naturally processed peptides from murine class II allele I-E^b (being phylogenetically related to HLA-DR) have shown that two to three preanchor residues are most common (6, 8). Therefore, we subjected self-peptide pools derived from five different HLA-D alleles to Edman sequencing to resolve these discrepancies.

The profiles obtained display allele- and isotype-specific signals of certain amino acids that are clearly favored. Those will be discussed here in detail.

Proline

Proline is the most abundant amino acid throughout the profiles of any allele (Fig. 1*A*), displaying two- or threefold the amount of other aliphatic residues. The signal maximum is located at position 2 (DR1, DR3, DR6, and DQw7) or at position 3 (DR5). Especially in the profiles of DR1, DR6, and DQw7 there is a tailing of proline in two to three cycles after the respective peak maximum. This is compatible with the hypothesis of shifting NH_2-termini. In addition to shallow shoulders appearing in the profiles of DR5 and DR6 around position 7, there are no further proline signals.

Phenylalanine

The aromatic side chain of phenylalanine appears to be important for binding of peptides to DR1 (position 2) as well as to DR6 (position 3) (Fig. 1*B*). In the latter case, the phenylalanine signals show significant tailing, comparable to that of proline in the DR6 profile, as cited above. No indication of phenylalanine is given for DR3, DR5, or DQw7.

Tyrosine

Tyrosine correlates with the signal of phenylalanine especially at position 2 of DR1 (Fig. 1*C*). DR6 shows a less pronounced signal at the same position. In addition, tyrosine peaks at position 4 of DR5 and is most prominent at position 5 of DR3.

Valine

Valine compares best to tyrosine in its allele-specific distribution (Fig. 1*D*): it has prominent maxima at position 2 of DR1 and DR6 as well as at position 4 of DR3. In contrast, DR5 displays two shallow maxima around positions 3 and 6.

Isoleucine

The profiles of DR6 (position 2) and DR3 (position 6) are dominated by signals from isoleucine (Fig. 1*E*). DR5 shows a plateau comprising positions 3 and 4, whereas isoleucine signals are lacking in the profile of DQw7.

Leucine

Leucine, similar to valine, is an amino acid that shows more or less pronounced peaks in the profiles of all four DR alleles but not in that of DQw7 (Fig. 1*F*): the most intensive signal is found at position 2 of DR5, tailing through positions 3 and 4, and at position 3 of DR1. DR6 shows a plateau comprising positions 2 and 3. Secondary maxima are found at positions 5 and 6 of DR5 and 6, respectively. Finally, DR3 displays a further maximum at position 8.

Lysine

Most striking, lysine carrying a charged side chain, shows significant signals comparable to leucine in its distribution (Fig. 1*G*), especially DR5 and DR6 at position 2 and DR3 to a smaller extent at position 6 contain lysine signals, whereas DR1 and DQw7 do not.

Glutamine

In this case, only in the profile of DQw7 is a prominent signal discernible (Fig. 1*H*) at position 2. Glutamine appears not to play a role in the peptides of any of the four DR alleles.

Asparagine

Asparagine mirrors the situation with glutamine (Fig. 1*I*): DQw7 shows a clear maximum at position 2, whereas all other alleles do not. Only DR3 displays a shallow plateau around position 9.

Aspartic acid

Aspartic acid is a typical representative of amino acids that do not show any reproducible signal, except for the DR3

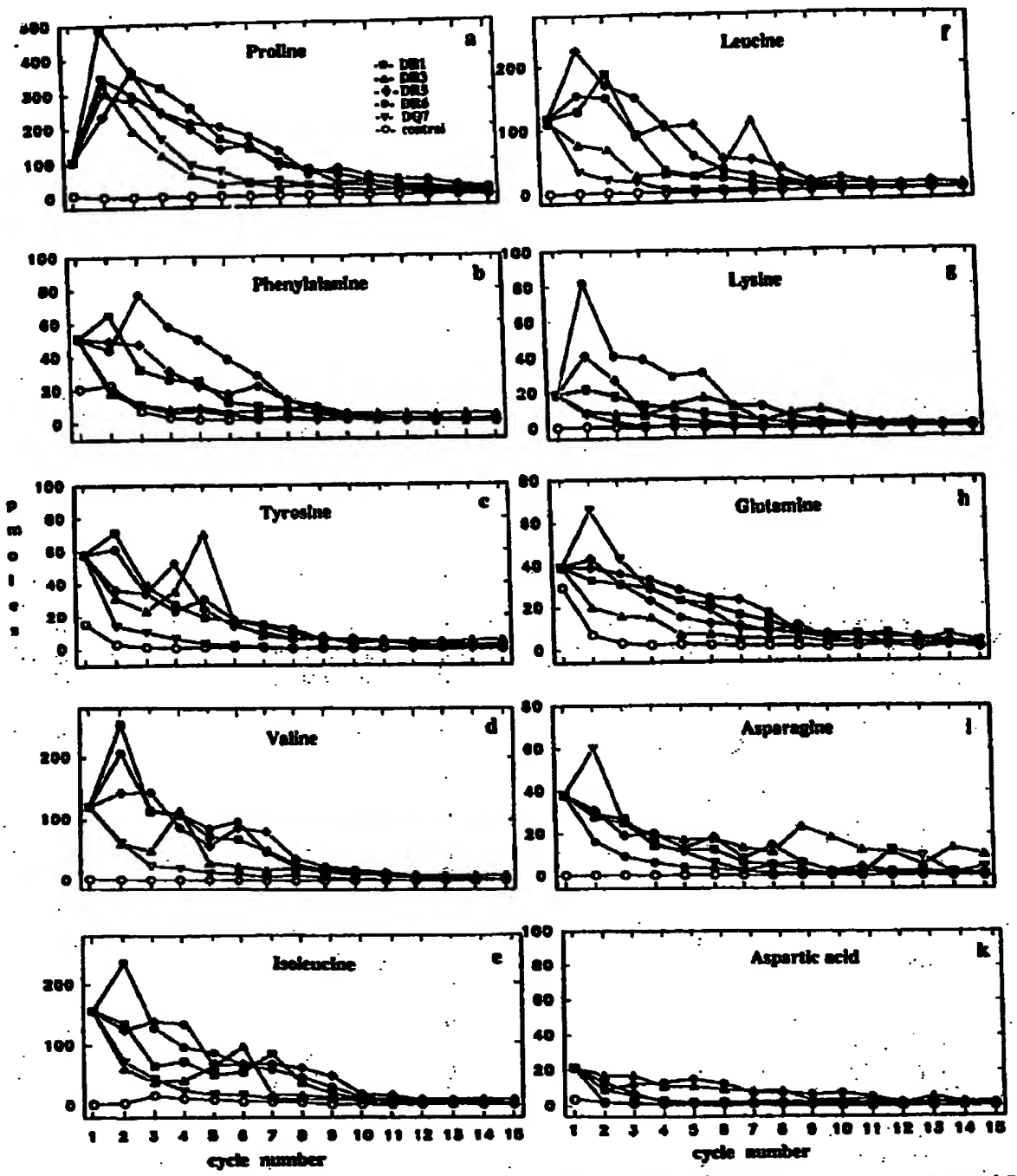

FIGURE 1. Sequencing profiles of naturally processed pooled peptides released from DR1, DR3, DR5, DR6, and DQw7. With the exception of aspartic acid (*K*), only amino acids with significant position-specific enrichment are shown. For clarity, the values of the first sequencing cycle of DR3, 5, 6, and DQw7 were normalized by the starting values of DR1. Material from a mock precipitate was sequenced as a control. Each bulk sequencing was performed in duplicate or triplicate.

profile where occasional signals around position 7 occurred. (Fig. 1*K*). It is the same with glutamic acid, arginine, histidine, serine, and threonine as well as with methionine, alanine, and glycine.

In vitro processing of cellular peptides by LAP

Because our self-peptide sequencing profiles revealed a significant accumulation of proline at positions 2 and 3

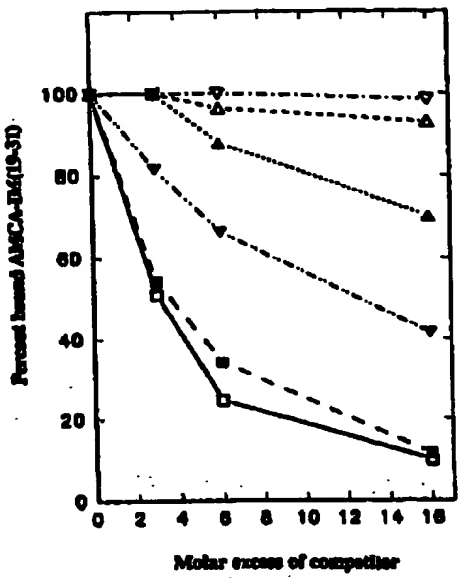

FIGURE 2. Competition assay by HPLC gel filtration with solubilized DR1 (3 μg), fluorescently labeled (AMCA)-IM(19–31) (0.1 μg), and different peptide fractions pretreated with LAP in vitro. The following peptides were investigated as competitors: cytosolic peptides, untreated (▽); LAP treated (▼); endosomal peptides, untreated (△); LAP treated (▲), naturally processed peptides released from DR1, untreated (□); LAP treated (■). LAP treatments were achieved as described in *Materials and Methods*. All measurements were performed in duplicate.

regardless of allele or isotype (Fig. 1*A*), we started to investigate, whether proline may be a remainder of NH_2-terminal processing. For this purpose, we prepared low-molecular-weight fractions isolated from the cytosol and endosomes of the EBV-B cell line WT-100 expressing DR1. From each preparation we separated the peptide fraction by reverse-phase HPLC. The highly complex mixtures that have been obtained were tested in a binding assay for their ability to compete with fluorescently labeled T cell epitope from influenza matrix protein, AMCA-IM(19–31). We have chosen IM(19–31), because it comprises the core region of the DR1-restricted epitope IM(17–31) carrying proline at position 2 (17).

As expected, crude untreated cytosolic peptides displayed no affinity for DR1 (Fig. 2): the binding signal of AMCA-IM(19–31) did not change on addition of cytosolic peptides. More striking, the same amount of endosomal peptides competed only slightly more. This result indicates that peptides occurring in the endosomal compartment are not preselected for binding to class II MHC molecules. In contrast, peptides that have been released from DR1 drastically reduced the binding of IM(19–31) owing to their high affinity.

In the second step we subjected all three peptide fractions to in vitro processing, exerted by LAP: LAP is known to cleave off amino acids from the NH_2-terminal rather unspecifically, but it cleaves Xaa-Lys, Xaa-Arg, Lys-Xaa, and Arg-Xaa bonds slowly and does not attack Xaa-Pro bonds at all (18). It turned out, that LAP treatment makes the affinity of cytosolic and endosomal peptides remarkably higher: competition rates of the cytosolic fraction were increased 30-fold by LAP treatment, whereas endosomal peptides reached only a fourfold increase of their original binding rates (Fig. 2). That means, both peptide fractions contained cryptic abilities to compete for binding to DR1. In contrast, naturally processed peptides from DR1 did not significantly alter their affinities for DR1 on LAP treatment. This finding corresponds to the sequencing profiles telling us that 30 to 50% of the DR1-associated peptides already possess a proline residue at position 2 or 3, precluding further NH_2-terminal trimming (Fig. 1*A*).

Aminopeptidase and cathepsin B treatment of DR1:self-peptide complexes

Previous studies suggested that class II MHC-associated self-peptides may protrude from the binding cleft at both ends, due to their lengths extending 15 residues (8, 11). According to this hypothesis, NH_2- as well as COOH-terminal peptide regions that are not protected by the MHC molecule may be further trimmed by exopeptidases. Therefore, we tested in an in vitro assay whether DR1-associated self-peptides could be further processed by treatment with AAP, which cleaves off NH_2-terminal amino acids rather unspecifically stopping at Pro-Xaa bonds (18), or cathepsin B, a cystein protease of the endocytic pathway that has been shown to play an important role in Ag processing (19, 20), probably owing to its well-established COOH-terminal peptidyl-dipeptidase activity. The effects we obtained on treatment of DR1-bound peptides are shown in the reverse-phase HPLC profiles in Figure 3: after coincubation with cathepsin B (Fig. 3*B*), there is a remarkable fraction of newly appearing peaks at the left side of the profile, corresponding to short retention times (20 to 50 min) compared with the untreated sample (Fig. 3*A*). At the same time, prominent peaks at the right side of the untreated mixture (65 to 85 min) have disappeared. Principally, the same has occurred when the fluorescently labeled peptide (AMCA)-DR1–33(6–14) was added as an internal control (Fig. 4): two new major peaks of fluorescence representing degradation products appeared with short retention times in the HPLC chromatogram (Fig. 4*C*), whereas the peak representing the intact peptide was lost (Fig. 4*A*). The shift to shorter retention times implies that a considerable portion of peptides has been further trimmed at the COOH-terminal. However, trimming at the NH_2-termini hardly occurred (Fig. 3*C*), the pattern of peaks has changed only marginally on AAP treatment, with no remarkable shifts being discernible. In contrast, the dose of AAP was sufficient to degrade the unbound control peptide DR1–33(5–16)-(AMCA), taking only 30 min (Fig. 4*D*). The corresponding

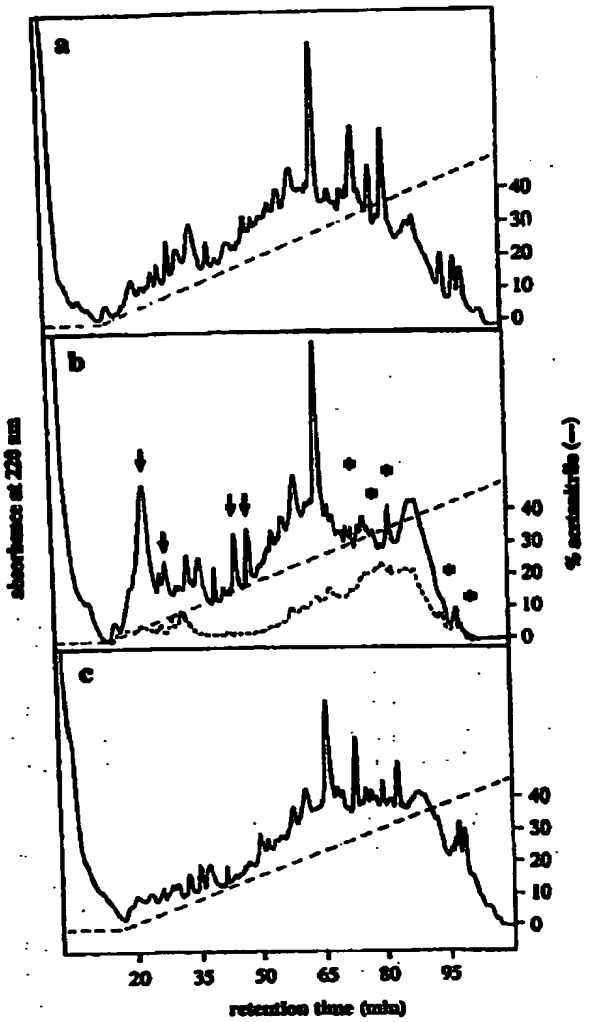

FIGURE 3. C18 reverse-phase HPLC profiles of DR1-associated peptides without protease treatment (*A*) and after treatment of DR1:self-peptide complexes with cathepsin B (*B*) or AAP (*C*). Peaks that have disappeared on cathepsin B treatment are marked by asterisks, newly generated peaks are marked by arrowheads. The hatched profile was obtained on cathepsin B treatment of DR1 molecules in absence of self-peptides.

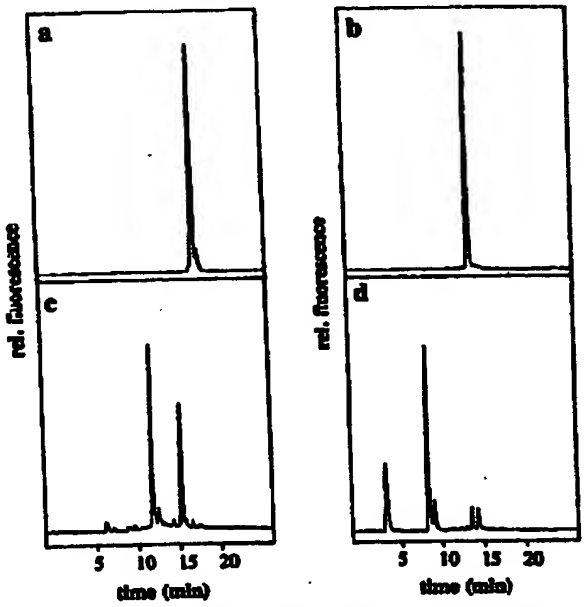

FIGURE 4. C4 reverse-phase HPLC chromatograms of (AMCA)-DR1-33(6–14) before (*A*) and after treatment with cathepsin B (*C*) and of DR1-33(6–16)-(AMCA) before (*B*) and after treatment with AAP (*D*). Digestions were conducted for 30 min in solutions containing DR1:self-peptide complexes (see Fig. 3). Peptides were eluted in a linear gradient of acetonitrile from 5 to 75% over a 30-min period.

bulk sequencing profiles of self-peptides after AAP treatment confirmed the protection of the peptides NH_2-termini from further trimming, because maxima of the postulated anchor amino acids near the NH_2-termini did not shift either (data not shown).

Discussion

The results presented herein extend previous sequencing data on naturally processed peptides bound to human MHC class II molecules (9, 10–12). To overcome the problem of fractionating mixtures of up to 2000 different peptides, as found with murine class II MHC I-A^d (7), we sequenced pooled peptides by Edman degradation. Our analyses focused on four HLA-DR alleles: DR1, DR3, DR5, and DR6. In addition, we investigated self-peptides from HLA-DQw7 as an isotype control.

Because of the heterogeneity in the site of terminal truncation (8, 11), pooled class II MHC-derived peptides proved to be more difficult to analyze compared with class I analogues. Nevertheless, we could deduce allele- and isotype-specific features (Fig. 1). Most characteristic, in the sequencing profile of each allele there is a limited number of amino acids that display increased occurrence at defined positions (Fig. 4). For example, tyrosine shows distinct signals comprising positions 2, 4/5, and 4 in the profiles of DR1, DR3, and DR5, respectively (Fig. 1*E*). In obvious contrast to the sharp maxima of single dominating amino acids at defined positions, termed "anchor" residues, which have been obtained from class I-associated self-peptide pools (3), the signals of class II self-peptides are less pronounced. An important reason is that the signal of a putative anchor residue is distributed to 1 to 3 sequential positions caused by the peptides' ragged NH_2-termini. Furthermore, our pool sequencing profiles reveal that anchor residues of similar physicochemical character as valine, leucine, or tyrosine most often are interchangeable, thereby contributing

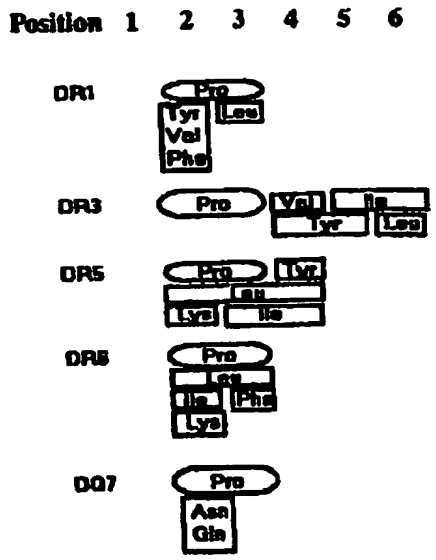

FIGURE 5. Putative anchor residues in the NH_2-terminal halves of naturally processed peptides from five investigated HLA-D alleles. The data are derived from the bulk sequencing profiles (see Fig. 1).

to the reduction of signal intensities, too. In view of the local degeneration of the signals we conclude that in most cases the number of residues preceding the first anchor residue is limited to 1 to 3. Otherwise, bulk sequencing profiles should not exhibit any maxima at all. Most of the self-peptide sequences of mouse class II MHC I-E^b (8) and of human DR11/Dw52 (12) proved to have NH_2-terminal presequences of limited lengths. However, sequences from various HLA-DR1-associated Ii peptides most recently published revealed presequences comprising 0 to 9 residues irrelevant for binding (11). This might be due to the fact that peptides of 20 to 25 residues may contain two agretopes in canonical arrangement, as previously shown (7). Alternatively, those Ii peptides may occupy a binding site that is different from the conventional peptide binding groove. The latter assumption is based on processing mutants transfected with DR3, which were found to display high amounts of the same Ii peptides without being able to convey SDS stability to DR3 molecules (21, 22).

Our sequencing data support previous approaches, whereafter mainly aromatic and aliphatic residues near the NH_2-terminal are crucial for anchoring peptides to HLA-DR molecules (9, 12, 15, 23, 24): e.g., in self-peptides from HLA-DR1 tyrosine, phenylalanine, and valine are highly abundant at position 2 and leucine peaks at position 3 (Fig. 5). In addition, at positions 2 and 3 there is a strong proline signal. That means, one hydrophobic amino acid or two sequential hydrophobic residues close to the NH_2-terminal are common features of DR1-associated peptides. This result is in agreement with the DR1 motif delineated by using an M13 epitope library (10); the first anchor was an aromatic residue (Tyr, Phe, or Trp), followed by an aliphatic one (Leu or Met) three residues downstream. Furthermore, the interaction of an immunodominant pertussis toxin peptide with DR1 proved to occur by Leu-33 (i) and Leu-36 (i + 3) (25). A further hydrophobic anchor, being spaced i,i + 8 or i,i + 9 relative to the NH_2-terminal one, has been found in various DR1-restricted foreign antigenic peptides (9, 24) by use of the M13 epitope library (10) and in a recently sequenced 16-mer self-peptide (9). In the bulk sequencing profiles of DR1 and of the other DR alleles there is no clear indication of further anchor residues comparable to those described for the region near the NH_2-terminal. We conclude that the bulky hydrophobic residue near the respective NH_2-terminal of DR1-associated peptides functions as a "primary" anchor absolutely necessary for binding, whereas further hydrophobic residues obeying spacings as i, i + 3 or i, i + 8 relative to the primary anchor should be classified as "secondary" anchors being optional but not essential for binding. This is consistent with the findings of Jardetzky et al. (26) suggesting a single hydrophobic residue near the NH_2-terminal to be sufficient for binding of the influenza hemagglutinin peptide HA(307–319) to HLA-DR1 (26). Preliminary results from the x-ray crystallographic analyses of DR1:HA(307–319) complexes confirmed the role of Tyr-309 as the primary anchor residue (T.S. Jardetzky, unpublished observations).

Furthermore, the definition of further anchor positions may be obscured by conformational variations caused by proline residues, as found with class I MHC HLA-Aw68; different degrees of bulging allow residues at positions 9, 10, or 11 to occupy pocket F (27).

DR1 pool sequencing profiles do not display a maximum for lysine (Fig. 1*F*), most recently suggested as a putative anchor residue in a comprehensive alignment of DR1-associated self- and foreign peptides (11). However, lysine is enriched in the profiles of other DR alleles, in DR5 and DR6 at position 2 together with leucine and isoleucine/leucine, respectively (Fig. 5). In the MHC class I field there is a similar example; sequence analyses with HLA-A3 revealed lysine together with tyrosine functioning as putative anchors at position 9 (28). In view of this co-occurrence of lysine with aliphatic or aromatic amino acids, the four methylene groups of the lysine side chain seem to be functional in making contact with a complementary hydrophobic pocket; the positively charged ϵ-amino group seems to be of less importance. In addition, lysine at position 2 may slow down the action of aminopeptidase. This fact may be relevant for peptides lacking proline at the NH_2-terminal (18).

Most probably, DR3, DR5, and DR6 share the same hydrophobic pocket with DR1, previously deduced from quenching studies (15) and most recently confirmed by fluorescence energy transfer (16). The enrichment near

the NH_2-terminal of chemically similar amino acid residues in naturally processed peptides of all investigated DR alleles strengthens this view (Fig. 5): it is isoleucine, valine, and tyrosine at positions 4 to 6 in DR3, tyrosine and leucine at positions 2 to 4 in DR5, phenylalanine and leucine at positions 2 and 3 in DR6. These findings are most easily explained by a hydrophobic pocket being formed by the monomorphic α-chain of any DR heterodimer.

Quite different from that, DQw7 displays asparagine and glutamine at position 2 emphasizing the isotype-specific nature of the postulated anchor residues. Glutamine and asparagine also proved to be part of the consensus motif of naturally processed peptides of murine I-A^b being closely related to HLA-DQw7 (8).

Facing the highly degenerate pattern of putative anchor residues relevant for binding to four different DR alleles, promiscuous binding of immunogeneic peptides to HLA-DR, as repeatedly found elsewhere (24, 29–31), seems quite natural. Nevertheless, our sequencing profiles also display substantial allelic differences. Most striking in this context, naturally processed peptides eluted from DR3 are distinguished from their allelic counterparts by the average lengths of preanchor sequences at their NH_2-termini (Fig. 5): compared with other DR alleles, the cluster of putative anchor residues of DR3 is shifted by at least two residues to the COOH-terminal. This observation correlates well with the Val-Gly dimorphism at position 86 of the DR β-chain. Gly-86-containing alleles as DR1Dw1 (as well as DR11Dw5 and DR14Dw16) have been previously shown to bind HA(307–319) by making contact with Tyr-309 at position 3, not tolerating a Ser-309 analogue (32). Val-86-containing alleles (as DR17Dw3), on the contrary, bound Ser-309-substituted analogues very well with the residue at position 5 becoming an important contact residue. This is in substantial agreement with isoleucine and tyrosine dominating at position 5 in naturally processed peptides of DR3 in our study (Fig. 5). More important, in a recently developed DR3-specific binding assay Tyr-7 at position 5 of the highly DR3-specific epitope Hsp65(3–13) turned out to be the primary anchor closest to the NH_2-terminal (33).

The sequencing data presented herein also have some relevance to the issue of natural processing of the peptides' NH_2-termini. Previous studies on murine I-E^b-associated peptides have suggested that cleavage by proteases of the cathepsin family is followed by trimming by exopeptidases (6). If it is true that class II processing is a stochastic event and that DR molecules have to select peptides from a complex mixture, as suggested most recently (11), amino acids functioning as stop signals for aminopeptidases may be found with high frequency at the NH_2-termini. That is in agreement with our results: proline that is known to terminate the proteolytic action of many different aminopeptidases (34), is enriched at positions 2 and 3 in any of the investigated alleles (Fig. 1*A*). Interestingly, proline at position 2 reappears in highly abundant Ii peptides isolated from DR1 (11) and a processing mutant expressing DR3 (21, 22). Nevertheless, Ii-derived peptides are not the only contributers of proline, because sequences of other single peptides eluted from DR1, DR2, and DR3 have revealed a more than 50% occurrence of proline at positions 2 and 3 (manuscript in preparation). For further investigation of this issue, we subjected endosomal and cytosolic peptides to LAP treatment. This processing simulation favors the view that the generation of a penultimate proline is sufficient to compete successfully with influenza peptide IM-(19–31) for binding to DR1. Because untreated cytosolic or endosomal peptides hardly showed any binding, shifting proline to position 2 by the action of LAP increases their affinity to DR1 considerably. The more pronounced effect of cytosolic peptides compared with their endosomal counterparts may be due to a more favorable length: endosomal peptides may have been more extensively processed by exo- or endoproteases confined to the endosomal compartment. As expected, naturally processed peptides revealed to be the most potent competitors regardless of a further LAP treatment. Their higher affinity must be ascribed to additional primary or secondary anchor residues at the proper positions. Because proline was not found to play a similar important role in DR1-restricted foreign peptides (30, 31), it appears to function as a secondary anchor besides preventing further NH_2-terminal trimming of peptides before or after binding to MHC class II molecules.

Provided that our model is correct, proline containing NH_2-termini should be accommodated in the binding groove making them inaccessible for attacking aminopeptidases. Therefore, we treated whole DR1:self-peptide complexes by AAP in vitro. There were no significant changes observed, either in the reverse-phase HPLC profile of the released self-peptide fraction (Fig. 3*C*) or in the bulk sequencing profiles (data not shown). Most peptides seemed to be protected from NH_2-terminal trimming. With proline occurring at positions 2 and 3 in only 40 to 50% of the pooled self-peptides (Fig. 1), not all of them are automatically protected from AAP action by proline. Therefore, steric hindrance by MHC molecules may contribute considerably to avoid destruction of the bound peptides' NH_2-termini.

In contrast, dramatic changes on cathepsin B treatment of DR1:self-peptide complexes point to considerable trimming of their COOH-termini (Fig. 3*B*). Because there is no indication of newly generated fragments when DR1 molecules freed from self-peptides by acid release were used (Fig. 3*B*), class II molecules appear to be highly resistant to cathepsin B treatment under the conditions described. We conclude that COOH-terminal portions of DR1-associated

self-peptides do protrude from the binding groove. This fact is in agreement with truncation variants from mouse class II I-E^b-associated self-peptides indicating COOH-terminal trimming in viv (6). Obviously, the transit time of DR: self-peptide complexes in endosomal or lysosomal compartments is not long enough to guarantee for complete COOH-terminal processing by carboxypeptidases like cathepsin B.

In conclusion, the analyses of pooled self-peptides from four DR alleles have revealed important structural features that are common to all of them: 1) there is an isotype-specific predominance of hydrophobic anchor residues with the distance from the NH_2-terminal significantly correlating with the DR β-chain dimorphism at position 86; 2) the length of NH_2-terminal preanchor sequences is limited to 1 to 3 residues for DR1, 5, 6 and 3 to 5 residues for DR3; and 3) proline functions as a stop-signal for N-terminal trimming and as a secondary anchor. This finding implies an important role for aminopeptidases during the processing of MHC class II-restricted epitopes. The highly degenerate occurrence of the described structural features rationalizes promiscuous binding of naturally processed peptides to allelic DR variants. Allele-specific restriction for T cell epitopes of foreign Ag may be due to peptide-specific combinations of primary and secondary anchor residues, sterical and conformational constraints by nonanchor residues, and the grade of processing at the NH_2- and COOH-termini. Because many different cellular aminopeptidases exert their proteolytic activity at the plasma membrane at neutral pH (33, 34), certain steps of the NH_2-terminal trimming of class II MHC-associated peptides may proceed after transit through the post-Golgi compartment, where loading is said to occur.

Note added in proof: In the meantime, the x-ray crystallographic analysis of HLA-DR1 clearly confirms the importance of an aromatic anchor residue close to the N-terminal for peptide-anchoring. (Brown, J. H., T. S. Jardetzkey, J. A. Gorga, L. J. Stern, R. J. Urban, J. L. Strominger, and D. A. Wiley. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature 364:33*).

Acknowledgments

We thank H.G. Rammensee, K. Falk, and O. Roetzschke for suggesting the sequencing of self-peptide mixtures and S. Stevanovic and G. Jung for peptide sequencing. We are especially grateful to A.-B. Vogt for useful discussions and carefully reading the manuscript.

References

1. Ziegler, H. K., and E. R. Unanue. 1981. Identification of a macrophage antigen processing event required for I-region restricted antigen presentation to T lymphocytes. *J. Immunol. 127:1869.*
2. Van Bleek, G. M., and S. G. Nathenson. 1990. Isolation of an endogenously processed immunodominant viral peptide from the class I H-2K^b molecule. *Nature 348:213.*
3. Falk, K., O. Roetzschke, S. Stevanovic, G. Jung, and H.-G. Rammensee. 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature 351:290.*
4. Jardetzky, T. S., W. S. Lane, R. A. Robinson, D. R. Madden, and D. C. Wiley. 1991. Identification of self-peptides bound to purified HLA-B27. *Nature 353:326.*
5. Hunt, D. F., R. A. Henderson, J. Shabanowitz, K. Sakaguchi, H. Michel, N. Sevilir, A. L. Cox, E. Appella, and V. H. Engelhard. 1992. Characterization of peptides bound to class I MHC molecule HLA-A2.1 by mass spectrometry. *Science 255:1261.*
6. Rudensky, A. Y., P. Preston-Hurlburt, H. Soon-Cheol, A. Barlow, and A. Janeway Jr. 1991. Sequence analysis of peptides bound to MHC class II molecules. *Nature 353:622.*
7. Hunt, D. F., H. Michel, T. A. Dickinson, J. Shabanowitz, A. L. Cox, K. Sakaguchi, E. Appella, H. M. Grey, and A. Sette. 1992. Peptides presented to the immune system by the murine class II major histocompatibility complex molecule I-A^d. *Science 256:1817.*
8. Rudensky, A. Y., P. Preston-Hurlburt, B. K. Al-Ramadi, J. Rothbard, and C. A. Janeway Jr. 1992. Truncation variants of peptides isolated from MHC class II molecules suggest sequence motifs. *Nature 359:429.*
9. Kropshofer, H., H. Max, C. A. Müller, F. Hesse, S. Stevanovic, G. Jung, and H. Kalbacher. 1992. Self-peptide released from class II HLA-DR1 exhibits a hydrophobic two-residue-contact motif. *J. Exp. Med. 175:1799.*
10. Hammer, J., B. Takacs, and F. Sinigaglia. 1992. Identification of a motif for HLA-DR1 binding peptides using M13 display libraries. *J. Exp. Med. 176:1007.*
11. Chicz, R. M., R. G. Urban, W. S. Lane, J. C. Gorga, L. J. Stern, D. A. A. Vignali, and J. L. Strominger. 1992. Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature 358:764.*
12. Newcomb, J., and P. Cresswell. 1993. Characterization of endogenous peptides bound to purified HLA-DR molecules and their absence from invariant chain-associated $\alpha\beta$ dimers. *J. Immunol. 150:499.*
13. Gorga, J. C., V. Horejsi, D. R. Johnson, R. Raghupathy, and J. L. Strominger. 1987. Purification and characterization of class II histocompatibility antigens from a homozygous human B cell line. *J. Biol. Chem. 262:16087.*
14. Ziegler, A., J. Heinig, C. Müller, H. Götz, F. P. Thinnes, B. Uchanska-Ziegler, and P. Wernet. 1986. Analysis by sequential immunoprecipitations of the specificities of the monoclonal antibodies TÜ22, 34, 36, 37, 39, 43, 58 and YD1/63.HLK directed against human class II antigens. *Immunobiology 171:77.*
15. Kropshofer, H., I. Bohlinger, H. Max, and H. Kalbacher. 1991. Self and foreign peptides interact with intact and disassembled MHC class II antigen HLA-DR via Trp pockets. *Biochemistry 30:9177.*
16. Kropshofer, H., H. Max, and H. Kalbacher. 1993. Evidence for cobinding of self- and allopeptides to human class II major histocompatibility antigen DR1 by energy transfer. *Proc. Natl. Acad. Sci. USA 90:403.*
17. Ceppellini, R., G. Frumento, G. B. Ferrara, R. Tosi, A. Chersi, and B. Pernis. 1989. Binding of labelled influenza matrix

peptide to HLA DR in living B lymphoid cells. *Nature 339:392.*

18. Beynon, R. J., and J. S. Bond, editors. 1989. *Proteolytic Enzymes—A Practical Approach.* IRL Press, Oxford, UK, pp. 233–238.
19. Guagliardi, L. E., B. Kappelman, J. S. Blum, M. S. Marks, P. Cresswell, and F. Brodsky. 1990. Co-localization of molecules involved in antigen processing and presentation in an early endocytic compartment. *Nature 343:133.*
20. Van Noort, J. M., J. Boon, A. C. M. van der Drift, J. P. A. Wagenaar, A. M. H. Boots, and C. J. P. Boog. 1991. Antigen processing by endosomal proteases determines which sites of sperm-whale myoglobin are eventually recognized by T cells. *Eur. J. Immunol. 21:1989.*
21. Riberdy, J. M., J. R. Newcomb, M. J. Surman, J. A. Barbosa, and P. Cresswell. 1992. HLA-DR molecules from an antigen-processing mutant cell line are associated with invariant chain peptides. *Nature 360:474.*
22. Sette, A., S. Ceman, R. T. Kubo, K. Sakaguchi, E. Appella, D. F. Hunt, T. A. Davis, H. Michel, J. Shabanowitz, R. Rudensdorf, H. M. Grey, and R. DeMars. 1992. Invariant chain peptides in most HLA-DR molecules of an antigen processing mutant. *Science 258:1801.*
23. Rothbard, J. B., and M. L. Gefter. 1991. Interactions between immunogenic peptides and MHC proteins. *Annu. Rev. Immunol. 9:527.*
24. O'Sullivan, D., T. Arrhenius, J. Sidney, M.-F. del Guercio, M. Albertson, M. Wall, C. Oseroff, S. Southwood, S. M. Colon, F. C. A. Gaeta, and A. Sette. 1991. On the interaction of promiscuous antigenic peptides with different DR alleles. Identification of common structural motifs. *J. Immunol. 147:2663.*
25. De Magistris, M. T., A. Di Tommaso, M. Domenghini, S. Chersi, A. Tagliabue, J. R. Oksenberg, L. Steinman, A. K. Judd, D. O'Sullivan, and R. Rappuoli. 1992. Interaction of the pertussis toxin peptide containing residues 30–42 with DR1 and the T cell receptors of 12 human T-cell clones. *Proc. Natl. Acad. Sci. USA 89:2990.*
26. Jardetzky, T. S., J. C. Gorga, R. Busch, J. Rothbard, J. L. Strominger, and D. C. Wiley. 1990. Peptide binding to HLA-DR1: a peptide with most residues substituted to alanine retains MHC binding. *EMBO J. 9:1797.*
27. Guo, H.-C., T. S. Jardetzky, T. P. J. Jardetzky, W. S. Lane, J. L. Strominger, and D. C. Wiley. 1992. Different length peptides bind to HLA-Aw68 similarly at their ends but bulge out in the middle. *Nature 360:364.*
28. DiBrino, M., K. C. Parker, J. Shiloach, M. Knierman, J. Lukszo, R. V. Turner, W. E. Biddison, and J. E. Coligan. 1993. Endogenous peptides bound to HLA-A3 possess a specific combination of anchor residues that permit identification of potential antigenic peptides. *Proc. Natl. Acad. Sci. USA 90:1508.*
29. Sinigaglia, F., M. Guttinger, J. Kilgus, D. M. Doran, H. Matile, H. Etlinger, A. Trzeciak, D. Gillessen, and J. R. L. Pink. 1988. A malaria T cell epitope recognized in association with most mouse and human MHC class II molecules. *Nature 336:778.*
30. Busch, R., G. Strang, K. Howland, and J. B. Rothbard. 1990. Degenerate binding of immunogenic peptides to HLA-DR proteins on B cell surfaces. *Int. Immunol. 2:443.*
31. O'Sullivan, D., J. Sidney, E. Appella, L. Walker, L. Phillips, S. M. Colon, G. Miles, R. Chesnut, and A. Sette. 1990. Characterization of the specificity of peptide binding to four DR haplotypes. *J. Immunol. 145:1799.*
32. Busch, R., C. M. Hill, J. D. Hayball, J. R. Lamb, and J. B. Rothbard. 1991. Effect of natural polymorphism at residue 86 of the HLA-DR β chain on peptide binding. *J. Immunol. 147:1292.*
33. Sidney, J., C. Oseroff, S. Southwood, M. Wall, G. Ishioka, F. Koning, and A. Sette. 1992. DRB1 *0301 molecules recognize a structural motif distinct from the one recognized by most DRβ$_1$ alleles. *J. Immunol. 149:2634.*
34. Mouritsen, S., M. Meldal, O. Werdelin, A. S. Hansen, and S. Buus. 1992. MHC molecules protect T cell epitopes against proteolytic destruction. *J. Immunol. 149:1987.*
35. Amoscato, A. A., A. Balasubramaniam, J. W. Alexander, and G. F. Babcock. 1988. Degradation of thymopentin by human lymphocytes: evidence for aminopeptidase activity. *Biochim. Biophys. Acta 955:164.*
36. Wu, Q., L. Li, M. D. Cooper, M. Pierres, and J. P. Gorvel. 1991. Aminopeptidase A activity of the murine B-lymphocyte differentiation antigen BP-1/6C3. *Proc. Natl. Acad. Sci. USA 88:676.*